com os dados do verbo *ser*), observar se efetivamente se está diante de um sujeito de 3ª pessoa plural, usando a análise das unidades entoacionais, se necessário, conforme se detalha no Capítulo 4 desta obra.

(3) Verificar se há não marcação de plural, com certo grau de produtividade, em formas verbais precedidas de consoante não nasal ou diante de pausa, contextos neutros em termos de influência fonética. Apenas em variedades em que se verifique a forma singular (não marcada) nesses contextos de forma produtiva (acima de 5%, segundo Labov 2003), pode-se falar em efetiva não marcação de plural.

(4) Assumir que a presença de forma aparentemente singular diante de consoante nasal, vogal oral ou vogal nasal não pode ser considerada por si mesma não marcação de pluralidade, devido ao provável efeito de sândi.

(5) Em se tratando de variedades cuja ocorrência de efetiva coincidência entre a forma de P3 e a forma de P6 em contextos neutros (sem qualquer motivação fonética) seja rara, buscar interpretações que a justifiquem, seja no plano morfossintático (Inacusatividade? Posposição do sujeito? Efeito cognitivo-processual da distância? Efeito semântico do traço inanimado? – cf. Vieira; Brandão, 2014), seja no plano fonético-fonológico (concordância "invisível"?; não ancoramento da nasal flutuante? – cf. Mota, 2013).

Os procedimentos aqui sugeridos visam a garantir o controle efetivo (i) das realizações fonéticas das marcas de P6 (o que se pode estender a outros contextos de expressão da concordância) e (ii) do contexto fonético e prosódico em que se inserem as ocorrências verbais.

Sem dúvida, para estabelecer o que seria o padrão natural, "não marcado", em cada caso, é necessário que se excluam da comparação entre as variedades as estruturas cuja variação dependa eminentemente de fatores morfossintáticos e fonéticos específicos, que, longe de expressarem um padrão particular, representam tendências gerais relativas a regras estruturais ou a processos fonético-fonológicos. É com esse propósito que se persegue, nos capítulos seguintes, a caracterização de variedades com base no levantamento dos elementos que permitam compreender a natureza da interface Fonética-Morfossintaxe nos padrões de concordância em Português.

2. A CONCORDÂNCIA DE TERCEIRA PESSOA DO PLURAL: PADRÕES EM VARIEDADES DO PORTUGUÊS

Silvia Rodrigues Vieira
Aline Maria Bazenga

INTRODUÇÃO[4]

Conforme se desenvolveu no Capítulo 1, a realização das marcas morfológicas de concordância tem sido um dos tópicos mais importantes para definir e caracterizar variedades do Português. No âmbito da formação do Português do Brasil (PB), o produtivo debate desenvolvido nos trabalhos científicos ancora-se em ao menos duas concepções distintas: uma que pressupõe que os índices brasileiros de cancelamento da marca de número tiveram origem no Português Europeu (PE) transplantado para o Brasil, de acordo com o movimento próprio da deriva linguística; e outra que propõe ser a falta de concordância brasileira resultante da intensa situação de contato linguístico, que teria dado origem a uma transmissão linguística irregular.

Sustentam a primeira concepção (Cf. NARO; SCHERRE, 2007) dados do Português arcaico, que demonstrariam a gênese da falta de concordância padrão, e do Português Europeu contemporâneo, que também evidenciariam a ausência da marca plural, embora com variações menos frequentes e com maiores limitações sociais. Os que defendem a segunda abordagem propõem, entretanto, que (i) a "variação da concordância no PE é na melhor das hipóteses residual e periférica" (LUCCHESI; BAXTER; SILVA, 2009, p. 331); e (ii) no PB, haveria um *continuum* relacionado à concretização da concordância – da norma urbana culta (com altos índices de concretização da marcação de plural) até as comunidades rurais afro-brasileiras (com os menores índices de marcação de plural) – que evidenciaria o efeito do contato entre línguas.

Como se pode observar, a área dos estudos linguísticos relativos à concordância enfrenta claramente um problema de validação empírica das propostas interpretativas existentes, visto estar relacionado sobretudo à

[4] Este capítulo constitui uma versão revista e traduzida do artigo intitulado "Patterns of third person verbal agreement", publicado originalmente no *Journal of Portuguese Linguistics*, v.12, n.2, em 2013.

comprovação científica das hipóteses o impasse entre as abordagens. Na realidade, os dados, não só em termos quantitativos, no que se refere à atestação da ausência de marcação de plural, mas também em termos qualitativos, no que se refere à natureza dos contextos de manifestação da marca, ora são tidos como pouco robustos, ora como suficientes. Ademais, soma-se à pouca descrição de dados do PE em materiais sociolinguisticamente controlados, o parco conhecimento das variedades africanas, que, por se constituírem em situação de intenso contato linguístico, poderão subsidiar também as interpretações das origens do PB.

Fica mais do que evidente, portanto, a necessidade de estudos que desenvolvam análises contrastivas das variedades do Português sob as mesmas condições e segundo os mesmos critérios. No que se refere à concordância de terceira pessoa plural (P6), Vieira (2011), com base em parte do material do banco de dados construído pelo Projeto de Cooperação Internacional, já referido nesta obra, que estava disponível à altura, apresentou resultados que permitiram atestar preliminarmente a existência de regras variáveis nas variedades brasileira e são-tomense do Português (PST), ao contrário do verificado na variedade lisboeta, em que se registrou uma inexpressiva produtividade da não marcação de plural associada a contextos estruturais bastante específicos (e por vezes até duvidosos), o que fez a autora supor o perfil de uma regra semicategórica de concordância. Bazenga (2012), valendo-se de critérios diferentes de coleta dos dados produzidos exclusivamente por falantes madeirenses, apresentou índices maiores de não realização da marca.

Na presente investigação, as autoras, em conjunto e adotando os mesmos critérios de seleção e tratamento dos dados, descrevem a totalidade das ocorrências de P6 na amostra básica do Projeto, a fim de dar conta dos seguintes objetivos: (i) estabelecer a produtividade das marcas de P6 nas variedades do Português europeias, brasileiras e, dentre as africanas, de São Tomé e Príncipe, definindo o estatuto de cada regra linguística em questão[5], consoante a tipologia proposta em Labov (2003); (ii) verificar a influência de fatores estruturais e sociais que atuem como restrições à regra variável da concordância, quando for o caso; e (iii) avaliar a relação entre fatores estruturais e sociais que condicionam o fenômeno e a caracterização do estatuto das variedades linguísticas do Português quanto aos padrões de concordância verbal.

[5] Aprofundando essa reflexão, Vieira; Brandão (2014) fundamentaram a postulação do estatuto das regras considerando não só a concordância verbal de P6, mas também a de primeira pessoa do plural (P4) e a concordância nominal.

Para cumprir os objetivos supracitados, estrutura-se o artigo da seguinte forma: apresentam-se, na Seção 1, informações gerais de cunho teórico-metodológico; na Seção 2, descrevem-se os resultados quanto à distribuição das ocorrências pela presença ou ausência da marca de número e quanto ao comportamento dos dados em relação às variáveis linguísticas e extralinguísticas controladas, detalhando as particularidades de cada variedade do Português e fazendo correlações entre os resultados, sempre que possível; e na seção final, discute-se o estatuto de cada variedade quanto ao fenômeno estudado, considerando as evidências quantitativas e qualitativas advindas da presente análise, contrastadas, quando possível, às de outras investigações.

1. FUNDAMENTOS SOCIOLINGUÍSTICOS, COLETA DE DADOS E DESCRIÇÃO DA *AMOSTRA*

Com base no aporte teórico-metodológico da Sociolinguística Variacionista, apresentado no Capítulo 1, o estudo fundamenta a análise da concordância, assumindo o princípio básico da heterogeneidade ordenada e da atuação de restrições estruturais e sociais sobre o fenômeno variável. Para a determinação do estatuto da regra linguística, a investigação adota a proposta de Labov (2003), segundo a qual ocorrem, além da regra variável, duas outras, a categórica e a semicategórica.

Como grupos de fatores a serem investigados, postula-se, por hipótese, a atuação das seguintes variáveis extralinguísticas relacionadas ao perfil do banco de dados estudado: faixa etária, sexo e escolaridade. De natureza estrutural, controlou-se a possível influência de variáveis que, com maior ou menor produtividade, já se mostraram relevantes em estudos anteriores: (i) quanto ao sujeito: posição em relação ao verbo, configuração morfossintática; distância entre o SN sujeito e o verbo; paralelismo entre as marcas do SN sujeito e as do verbo (paralelismo oracional); traço semântico (animacidade); e (ii) quanto às formas verbais: grau de diferenciação fônica entre as formas singular e plural (saliência fônica); transitividade; paralelismo entre as marcas dos verbos em série (paralelismo discursivo).

Para a coleta de dados, determinou-se, de forma rigorosa, o tipo de ocorrência que seria considerada efetivamente variável, aquela que remete semântica e formalmente a P6, nas mais variadas expressões de sintagmas

nominais, estando eles explícitos ou subentendidos. Por essa razão, qualquer contexto que acarretasse interpretação ambígua (MARTINS, 2012) quanto à referência plural do sujeito ou da forma verbal foi desconsiderado na análise quantitativa dos dados, de modo a garantir a comparabilidade dos resultados em todas as variedades.

Os critérios para coleta dos dados atendem ao pressuposto de que um enunciado pode remeter a estruturas sintáticas diferentes, o que implica assumir que nem sempre a aparente falta de expressão mórfica de plural constitui efetivamente um caso de não concordância. Nesse sentido, não compõem o conjunto de dados as seguintes construções: formas verbais de P6 empregadas como recurso de indeterminação (*Quebraram a janela. Não sabemos quem fez isso.*); aparentes sujeitos de terceira pessoa plural (expresso ou nulo) que remetem a constituintes topicalizados (*Pães de mel, (isso) é meu lanche favorito*); construções com verbo copulativo, que podem ser interpretadas como apresentacionais, com ou sem clivagem (*os meus filhos é/são o meu problema; são/é eles que me incomodam*); formas verbais infinitivas, dada a frequente ambiguidade na referência, se pessoal ou impessoal; ocorrências de *ter* e *vir* no presente do indicativo, por as formas singular e plural serem homófonas ou de difícil reconhecimento na maior parte dos dados do *corpus*; outras construções cuja determinação da forma verbal como singular ou plural é de definição duvidosa por motivos fonético-fonológicos (como na adjacência da desinência plural seguida de palavra com consoante nasal ou com vogal, como em, respectivamente, *eles cantam na praça* e *eles cantam uma canção calma*), além das diversas construções com coordenação, expressões partitivas, numéricas, títulos de obras, dentre outras, cuja interpretação possa ser singular ou plural.

Os dados foram coletados de todas as entrevistas que constituem a amostra básica do *Corpus* Concordância[6], conforme apresentação no primeiro capítulo deste livro, aqui retomada. Em síntese, trata-se de 18 informantes em cada ponto de inquérito, distribuídos por sexo, três faixas etárias (18-35 anos; 36-55 anos; 56-75 anos) e três níveis de escolaridade (fundamental: 5 a 8 anos; médio: 9 a 11 anos; superior: mais de 11 anos). Para as variedades europeias, consideram-se os pontos referentes a Oeiras (OEI) e

[6] Colaboraram na tarefa de coleta e tratamento dos dados de concordância verbal de 3ª pessoa plural nas variedades brasileiras os alunos Juliana Cristina de Paula Pires, Vitor de Araujo Ferreira, Cristina Marcia Monteiro de Lima Correa, Fernanda Villares Vianna Barreto, Rodrigo Cunha da Silva e Daiane Rangel da Silva, e na variedade são-tomense a graduanda Karen Cristina da Silva.

Cacém (CAC), na Região Metropolitana de Lisboa, e de Funchal (FNC), na Ilha da Madeira; para as variedades brasileiras, os pontos referentes a Copacabana (COP) e Nova Iguaçu (NIG), na Região Metropolitana do Rio de Janeiro, e para a variedade do Português falado em São Tomé e Príncipe, tem-se a amostra relativa a São Tomé (PST). Para a variedade são-tomense, cumpre esclarecer que se considerou adicionalmente, em uma rodada com a finalidade específica de averiguar o efeito da frequência de uso da língua materna, a totalidade das entrevistas da amostra são-tomense que se enquadram nos critérios da presente investigação.

A seguir, expõe-se a distribuição das 7063 ocorrências coletadas neste estudo pelas variedades do Português:

Tabela 1. Número de dados de formas verbais de P6 por variedade

PE			PB		PST
OEI	CAC	FNC	COP	NIG	ST
1.467	1.185	914	1.395	1.365	737
	3566		2.760		737

2. A CONCORDÂNCIA DE P6 EM VARIEDADES DO PORTUGUÊS

O tratamento dos dados demonstrou que a expressão da concordância ocorre segundo padrões diferenciados nas variedades do Português. A tabela a seguir confirma essas diferenças.

Tabela 2. Distribuição dos dados com e sem marca de concordância verbal de P6

	SVs de terceira pessoa plural			
Amostra	Com marca de número		Sem marca(s) de número	
	Ocorrências	**Percentual**	**Ocorrências**	**Percentual**
Oeiras	1454/1467	99,1	13/1467	0.9
Cacém	1176/1185	99,2	9/1185	0,8
Funchal	866/914	94,7	48/914	5,3
São Tomé	679/737	92,1	58/737	7,9
Copacabana	1229/1395	88,1	166/1395	11,9
Nova Iguaçu	1067/1365	78,2	298/1365	21,8

Os índices gerais permitem afirmar que as amostras do PE, grosso modo, apresentam uma regra semicategórica de concordância de terceira pessoa, embora Funchal demonstre comportamento no limite entre o semicategórico e o variável, nos termos de Labov (2003). De outro lado, os índices de marcação de plural nas demais variedades urbanas, tanto no PST quanto no PB, sinalizam o comportamento de uma regra variável, com forte preferência pela concordância.

Obviamente, esses índices brutos, considerando todos os contextos, não permitem ver a complexidade do comportamento dos dados, o que exige uma análise da atuação das variáveis independentes que se mostraram relevantes no tratamento estatístico realizado pelo Goldvarb-X. Nas amostras do PE, continental (Oeiras e Cacém) ou insular (Funchal), não foi possível estabelecer a lista de grupos de fatores relevantes a partir da seleção estatística, dada a limitada concretização da não concordância. Desse modo, parte-se da análise detalhada das construções em que não se concretizou a marca de plural e, especificamente em Funchal, de alguns índices percentuais.

2.1. PADRÕES DE CONCORDÂNCIA EM PE: UMA REGRA VARIÁVEL?

No PE, conforme já se observou, os resultados globais conduzem a postular a regra da concordância verbal como semicategórica. O Quadro 1[7] apresenta todas as ocorrências de cancelamento da marca de plural, num total de 22 contextos nas amostras das variedades continentais.

[7] Em todo o livro, os exemplos extraídos do Banco de dados do Projeto são identificados da seguinte forma: CAC = Cacém; OEI = Oeiras/Lisboa; A = 18 a 35 anos; B = 36 a 55 anos; C = 56 anos em diante; 1 = até 5 anos de escolaridade; 2 = 6 a 9 anos de escolaridade; 3 = 10 a 12 anos de escolaridade; H = homem; M = mulher.

Quadro 1. Controle de todas as ocorrências de P6 sem marcação de plural no PE

CACÉM/ OEIRAS	Ocorrências	FATORES LINGUÍSTICOS				
		I	II	III	IV	V
		Sândi Externo	VS	QUE	[-anim]	T. Verbo
CACA3M	(1) e depois a partir da estação **existe** os autocarros para os diversos sítios	x	x		x	x inac **existir**
CACA3M	(2) não existe cinema não existe teatro não **existe**... exposições esse tipo de não existe	x	x		x	x inac **existir**
CACC1H	(3) chegou ao fim do mês não pagou aquela/quinhentos euros só pagou trezentos **fica** os outros duzentos a pagar	x	x		x	x cop
CACC3H	(4) **existe** vários tipos de população relativamente à classificação que deveremos desenvolver aplicar		x		x	xinac **existir**
CACA1M	(5) tem umas feiçõezinhas que **identifica** logo que elas são brasileiras.			x		
CACB1M	(6) vai ver a bola... claro juntamente com outros amigos que se **junta** a ver a bola..	x		x		
CACB2M	(7) oitenta mil escudos **equivale** neste momento a quatrocentos euros...	x			x	
CACB3M	(8) é assim a de antigamente tinha as suas partes chatas que **era** aquelas regras completamente absurdas	x	x	x	x	x cop
CACB3M	(9) os meus pais não nos deixavam ir pra rua brincar não nos **deixava** se calhar receber amigos em casa					
OEIA3M	(10) **saiu** algumas gramáticas alteradas mas ninguém está a usar nada	x	x		x	x inac
OEIB1M	(11) porque eu já sabia de quem era não sabia de quem **era** os carros mas sei sei que é dos vizinhos	x	x	x	x	x cop
OEIB1M	(12) um bocado a situação que **aparece** pessoas que a gente pensa às vezes que nem que não precisam		x			x inac

OEIC1M	(13) portanto **continua** os tempos a evoluir cada vez mais não é?	x	x		x	x cop
OEIB2M	(14) ah ah há *coisas* que noto que **leva** muito tempo para se conseguir uma consulta hum hum	x		x	x	
OEIC1H	(15) os hábitos que não me **agrada** muito a é é é - neste momento que estamos a atravessar é a falta de respeito	x		x	x	
OEIC2H	(16) pelo menos por aquilo que me dizem e por aquilo que vejo ah - *coisas* que- no meu tempo - **era** muito diferente	x		x	x	x cop
OEIC3M	(17) e eles são cumpridores (...) quando têm que estudar não- primeiro estudam - o que é bom também (D-isso é óptimo) nunca **foge** nem pouco mais ou menos não é?	x				x intr
OEIC3M	(18) os livros dantes eram um bocado agora agora vejo a diferença acho achava que **era** assim era um bocado massudos não é?				x	x cop
OEIB3M	(19) acho que antigamente as pessoas tinham mais mais tempo mais tempo porque **tava** mais em casa	x				x cop
OEIC1M	(20) para esses há crise e também aquelas pessoas que vieram de fora também **chegou** aqui também não têm nada não é?	x				x inac
OEIA3H	(21) estes óculos estão a estão a cinquenta euros mas querem aumentá-los para sessenta e cinco - então metem o preço em cima sessenta e cinco euros e depois metem riscado os os cinquenta euros tipo a dizer que **está** em promoção	x			x	x cop
OEIB1M	(22) as coisas que a gente conhece também não *é coisas* assim é os carros –	x	x		x	x cop
Totais		17	12	7	15	16

Nas raríssimas ocorrências sem marca de número nas amostras de Cacém e Oeiras, 9 e 13 respectivamente, verificam-se sobretudo contextos marcados, em que estão presentes estruturas que, ao que parece, universalmente[8] favorecem a não realização da marca, a saber: sândi externo (coluna **I**), quando o verbo é seguido por [+vogal] ou [cons.+nasal] no início da palavra seguinte, em que ocorre a não ancoragem do autosegmento /N/ ao nó da Rima, na estrutura silábica (MOTA; MIGUEL; MENDES, 2012, p. 171), em 17/22 ocorrências (72,2%) fenômeno que afeta principalmente verbos de menor saliência fônica; sujeito posposto (coluna **II**), em 12/22 ocorrências (54,5%); a presença do relativo QUE (coluna **III**), em 7/22 dados (38,8%); sujeito inanimado (coluna **IV**), com 15/22 (68,1%); e tipo de verbo (coluna **V**), em particular, os verbos intransitivos, inacusativos e copulativos (CARDOSO; CARRILHO; PEREIRA, 2011), num total de 16/22 (72,7%). Para além desses elementos linguísticos, destacam-se os exemplos com sujeito anteposto onde se constata a provável intervenção de fatores de ordem cognitiva, "inerentes à produção do discurso oral", relacionados com a "planificação do discurso e a quantidade de material linguístico entre o sujeito e o verbo" (MOTA; MIGUEL; MENDES, 2012, p. 169).

A aplicação da regra de concordância verbal no âmbito das variedades do PE conduz, no entanto, a uma polaridade entre as variedades continentais e a variedade insular, as primeiras situando-se mais perto da sua aplicação categórica e a segunda no limiar da regra variável. Essa polaridade interna à variedade europeia é confirmada pelos índices relativos à atuação dos fatores de natureza linguística, sobretudo os de natureza sintático-semântica, sendo mais tênue quanto ao papel desempenhado pelo *Traço semântico* do sujeito frásico (Tabela 3) e mais acentuada, quando está em causa a atuação da variável *Posição do sujeito* (Tabela 4).

[8] A esse respeito, vale conferir Brandão; Vieira (2012b, p. 1061), que apresentam evidências fundamentadas em Corbett (2000, p. 201), retomadas ao fim do presente artigo, para sugerir a universalidade de algumas variáveis estruturais que atuam na aplicação da concordância verbal.

Tabela 3. Atuação da variável *Traço semântico do Sujeito* para a implementação da marca verbal de P6 no PE

Variedade / Amostra		Sujeito [+animado]	Sujeito [-animado]
P E	OEI	1249/1253 99,7%	205/214 95,8%
	CAC	970/972 99,8%	206/213 96,7%
	FNC	728/746 97,6%	138/168 82,1%

Tabela 4. Atuação da variável *Posição do Sujeito* para a implementação da marca verbal de P6 no PE

Variedade / Amostra		Sujeito anteposto	Sujeito posposto
P E	OEI	777/782 99,4%	36/40 90,0%
	CAC	731/735 99,5%	40/44 90,9%
	FNC	550/573 96,0%	18/34 47,2%

Nessa comparação, surge como particularmente relevante a tendência observada nos dados de Funchal para o cancelamento da marca do plural em contexto de sujeito posposto (47,2%). Nesse contexto, observou-se que, na amostra de Cacém (Quadro 1), chamam a atenção as três ocorrências do item verbal inacusativo *existir* sem marca de plural (exemplos (1), (2) e (4)), em articulação com a posição do sujeito (posposto, coluna II) e sua natureza semântica (traço [-animado], na coluna IV). Esse fenômeno tem maior difusão na variedade insular. Atendendo unicamente ao contexto de sujeitos pospostos, registra-se um total de 20 ocorrências de *existir* em 34, ou seja 58,8% (11 com marca de plural e 9 sem marca). Do conjunto das 18 ocorrências sem marca de plural, esse item seria, assim, responsável em 50% pelo cancelamento da marca do plural nessa variedade.

O condicionamento morfofonológico em contexto de sândi externo – que se traduz pela redução do ditongo nasal decrescente da desinência de P6, motivada pelo contexto fonético à direita do verbo – atua, ao que parece, na maioria dos dados das duas variedades continentais (Quadro 1, coluna I) e, de forma ainda mais expressiva, na variedade insular: em 68,7% dos casos em que se observa a ausência de marca do plural, conforme ilustrado em (1),

para os contextos em que estão presentes verbos com baixa saliência fônica (com pouca diferenciação entre as formas singular e plural), independentemente de outras configurações sintáticas:

(1) a. seguido de vogal (21 dados)
as mercearias na altura fechava às onze (FNCB1M)
b. seguida de nasal (7 dados)
houve crianças as que tinha mais (FNCB1M)
c. seguida de pausa (2 dados)
os cá de fora que se lixe (FNCC3M)

Tal como nas outras chamadas variedades dialetais do PE continental (MOTA; RODRIGUES; SOALHEIRO, 2003; MOTA; MIGUEL; MENDES, 2012; RODRIGUES, 2012), ocorrem, na variedade urbana insular, variantes flexionais não-padrão. Essas variantes não-padrão são de dois tipos:

i. variante representada por –EM, do ditongo nasal [ɐ̃j̃] de P6 de verbos com vogal temática /e/ e /i/ aos verbos e que ocorrem em verbos com vogal temática /a/, padrão que reflete a convergência ou nivelamento de variantes;

ii. variante representada por –U, que pode corresponder a vogal nasal [ũ] ou oral [u], dependendo o traço nasal do contexto fonético seguinte (vogal ou consoante nasal).

Esses dois padrões representam cerca de 9% dos dados, ou seja, 83 em 866 ocorrências totais de marcação da concordância verbal, e quase o dobro das realizações sem a marca padrão do plural de 3ª pessoa (5,3 %).

A realização [ɐ̃j̃] de P6 padrão, presente no paradigma do presente do indicativo dos verbos com VT /e/ e /i/, estende-se aos verbos com VT /a/, estabelecendo uma convergência na marcação de P6. Este processo de nivelamento na marcação de P6 é também observado nos paradigmas do pretérito imperfeito do indicativo e do pretérito perfeito indicativo, substituindo P6 em [ɐ̃w̃] padrão nestes paradigmas (Tabela 5).

Tabela 5. Número de ocorrências de variantes P6 não-padrão em diferentes paradigmas verbais (Funchal)

Variantes P6 não-padrão	Presente	Pretérito Imperfeito			Pretérito Perfeito			Totais
		Ind.			Ind.			
	Ind.	VT /a/	VT /e/	VT /i/	VT /a/	VT /e/	VT /i/	
-EM	19	2	16	9		-	3	75
-U	-	5	1	2		-	-	8

A realização da variante não-padrão com vogal oral [u] ou nasal [ũ] está circunscrita ao pretérito imperfeito do indicativo. Os paradigmas afetados correspondem, em sua maioria, aos contextos de baixa saliência fônica (diferença circunscrita ao acréscimo de marca de nasalidade, com alteração da qualidade vocálica ou não), exceção feita a oito ocorrências (uma do tipo *vaiem-se* por *vão* e sete correspondentes aos contrastes de terceira pessoa do Pretérito Perfeito). Na variedade de Funchal, verificam-se, assim, padrões de variantes de flexão de P6 em concorrência e que seriam, no caso do imperfeito do indicativo, os seguintes: ditongos decrescentes nasais [ɐ̃w̃] padrão e [ɐ̃j̃] não-padrão; sem ditongo nasal, com as realizações não-padrão, [ũ] e [u], ilustradas pelos exemplos em (2) – (4).

(2) Variante sem marca (variante sem ditongo, P3=P6)
a. seguida de vogal:
elas à que chegava à escola (FNCB1M)
b. seguida de nasal:
[problemas] que penaliza mais os seus clientes (FNCC3H)
c. seguida de pausa:
os cá de fora que se lixe (FNCC2M)
d. seguida de consoante não-nasal:
mas os dias foi [foram] passando (FNCA1M)
e. verbo *existir:*
mas existe tarifários melhores (FNCA3M)

(3) Variante – EM "convergente"
a. *mas já* ***trabalharem*** *aqui nestas zonas todas* (FNCC1H)
b. *pessoas a trabalhar comigo que* ***tinhem*** *vergonha* (FNCC2M)
c. *elas que* ***vaiem-se*** *[vão-se] matando* (FNCA1M)

(4) Variante – U "conservadora"
a. *eles* ***vinho*** *brincar* (FNCB1M)
b. *os meus avós* ***ero*** *padrinhos da minha irmã Gorete* (FNCB1M)

No caso das variantes sem ditongo, coloca-se a questão de estar em jogo um processo similar ao descrito para a variedade de Braga (RODRIGUES, 2012, p. 221), em que, contrariamente ao padrão, /N/ é associado ao núcleo. As formas [o] e [u] resultariam de um processo de desnasalização do ditongo nasal em [w] e posterior redução do ditongo nas vogais labiais. A forma [ũ] corresponderia apenas a um processo de redução do ditongo mas sem a ocorrência de desnasalização.

A verificar-se esta possibilidade no âmbito da variedade insular de Funchal, esta variante poderia ser considerada como "conservadora", resultante da difusão de alguns aspectos observados em subvariedades dos dialetos setentrionais do PE, geralmente caracterizados por traços que não sofreram algumas das mudanças ocorridas no dialeto que viria a constituir o PE padrão.

No caso das outras variantes, sem ditongo, por um lado, e com ditongo de convergência, em verbos com VT /a/, /e/ e /i/ do presente do indicativo (e que se estende a outros paradigmas verbais), por outro, poderá estar em causa apenas a associação ou não do autossegmento flutuante /N/, tal como a representação formulada para o PE padrão por Mateus; Andrade (2000, p. 133). Segundo esta proposta, as variantes sem ditongo podem resultar da propriedade de leveza que caracteriza os ditongos nasais finais não acentuados, de 3ª pessoa do plural em formas verbais do PE padrão. Estes ditongos, mas também aqueles que ocorrem em formas nominais simples (*homem*) ou com sufixo *-agem* (*paragem, lavagem*)[9], são considerados pós-lexicais pelo fato de se encontrarem em palavras marcadas pela ausência de constituinte temático (por oposição aos ditongos nasais lexicais, gerados no léxico e que atraem o acento para o final de sílaba (*irmão*) e que admitem, como único segmento em coda, a fricativa /S/ (*irmãos*). Nos ditongos pós-lexicais, a semivogal é epentética, atendendo

[9] A redução de ditongos nasais decrescentes deste tipo também ocorre na variedade insular, mas com menor frequência do que observado em variedades do sul do Brasil (SCHWINDT; SILVA; QUADROS, 2012, p. 353). Nessas variedades, a classe lexical parece atuar na redução de nasalidade, sendo as formas nominais mais sujeitas à redução do que as formas verbais de P6.

à inexistência de uma vogal marcadora de classe temática em nível lexical, e surge após o processo de ditongação, os dois segmentos ocupando uma única posição no núcleo. O autossegmento flutuante /N/ projeta-se sobre o núcleo silábico, nasalizando os dois segmentos em simultâneo[10].

No que diz respeito às variáveis extralinguísticas, tanto a variável escolaridade (Tabela 6) como a variável sexo (Tabela 7) têm maior incidência na variedade insular.

Tabela 6. Atuação da variável *Escolaridade* para a implementação da marca verbal de P6 no fenômeno PE

Variedade / Amostra		Nível fundamental (5 a 8 anos)	Nível médio (9 a 11 anos)	Nível superior (12 a 15 anos)
P E	OEI	411/417 98,6%	386/388 99,5%	657/662 99,2%
	CAC	374/377 99,2%	390/391 99,7%	412/417 98,8%
	FNC	212/236 89,8%	257/268 95,9%	397/410 96,8%

Tabela 7. Atuação da variável *Sexo* para a implementação da marca verbal de P6 no PE

Variedade /Amostra		Masculino	Feminino
P E	OEI	757/760 99,6%	697/707 98,6%
	CAC	523/525 99,6%	653/660 98,9%
	FNC	454/472 96,2%	412/442 93,2%

[10] Por outro lado, talvez seja de considerar, para além da leveza desse tipo de ditongos, um outro tipo de instabilidade que afeta os ditongos orais quando seguidos de fricativa em coda, tais como [ajʃ] > [aʃ] (mais > mas) e [ojʃ] > [oʃ] (depois > depos) (BOLEO, 1975, p. 338) e também característico de outras variedades dialetais centro-meridionais do PE e insulares (Ilha Terceira do Arquipélago dos Açores, e no sul do Brasil (BRESCANCINI, 2009). De referir ainda um outro fenômeno fonético que consiste na redução do [u] átono final (CINTRA, 2008 [1990], p. 181) perceptível como [ɨ] / ou sem realização, e que afeta qualquer palavra que o contenha, independentemente da sua classe lexical. No que diz respeito aos verbos, as formas de primeira pessoa do singular e do plural do presente do indicativo são realizadas sem a vogal [u], (*falo > fale; tenho > tenhe; falamos>falames; temos> temes).*

Os efeitos estatísticos da atuação das variáveis extralinguísticas[11] na amostra de Funchal singularizam esta variedade no âmbito dos dados do PE. Uma observação mais pormenorizada dos dados mostra que os informantes com baixa escolaridade são responsáveis por maiores índices de cancelamento de marca de plural e também por maiores valores de produção de variantes não-padrão de P6, em contextos estruturais caracterizados por enunciados com sujeitos antepostos e sujeitos não expressos (Tabela 8).

Tabela 8. Atuação da variável *Escolaridade* no cancelamento de marca e produção de marcas não-padrão de P6 em contexto de sujeitos antepostos e sujeitos não expressos

	Com marca não-padrão		Sem marca	
	Ocorrências	%	Ocorrências	%
Nível fundamental (5 a 8 anos)	51/82	62,2	20/30	66,6
Nível médio (9 a 11 anos)	31/82	37,8	5/30	16,7
Nível superior (12 a 15 anos)	-		5/30	16,7

Quando considerados os valores totais da produção de variantes flexionais de terceira pessoa do plural não-padrão e de variantes homófonas de terceira pessoa do singular (resultantes de cancelamento de marca de concordância verbal), a atuação da variável *sexo* é bastante evidente. As mulheres são responsáveis por estratégias não conformes ao PE padrão (Tabela 9).

Tabela 9. Atuação da variável *Sexo* no cancelamento de marca de P6 e na produção de variantes com marca não-padrão (-EM e -U)

	Com marca não padrão		Sem marca	
	Ocorrências	%	Ocorrências	%
Mulheres	65/82	79, 3	30/48	62,5
Homens	17/82	20, 7	18/48	37,5

[11] O índice de 5,3% de não marcação de número na amostra do Funchal permitiu fazer rodadas variacionistas, cujos resultados apontam a atuação destas duas variáveis extralinguísticas, respectivamente em segundo e quarto lugar, no conjunto de variáveis ordenadas por relevância estatística e respectivos pesos relativos: *posição do sujeito, escolaridade, animacidade, sexo* e *paralelismo discursivo.*

Esses valores permitem considerar, no seio da variedade urbana insular, a existência de "uma situação sociolinguisticamente polarizada, reforçada pelo fato de o nível de escolaridade (...) se associar a outras variáveis, como nível socioeconômico, maior ou menor acesso a bens culturais, maior ou menor contato com indivíduos de diferentes estratos sociais" (BRANDÃO; VIEIRA, 2012b, p. 1054). Parece claro existir no seio da sociedade madeirense uma nítida fronteira que separa os falantes pouco escolarizados, estigmatizados pela sua linguagem popular[12] e as elites que se manifestam pelo uso consolidado do PE padrão. A cidade de Funchal tem sido ao longo do tempo polo de atração de migrações internas, nela convergindo população rural, que se concentra na sua periferia e contribui para a atividade econômica ligada aos serviços e comércio (Censos 2001 e 2011).

Contrariamente à generalização formulada por Labov (1990, p. 205), segundo a qual as mulheres, no mesmo grupo social, tenderiam a ser mais conservadoras no que se refere à preferência pelas variantes padrão em relação às estigmatizadas (não-padrão), a amostra de Funchal (Tabela 10) mostra precisamente o contrário: na comunidade de falantes com baixa escolaridade, as mulheres observam comportamentos linguísticos marginais, liderando o uso de variantes com marca não-padrão (vogal –U e com ditongo nasal –EM) com a porcentagem de 69,2% e, com menor saliência, o de variantes sem marca de pluralidade (56%).

Tabela 10. Atuação da variável *Sexo* no cancelamento de marca de P6 e na produção de variantes com marca não-padrão na fala de indivíduos com nível fundamental de escolaridade

Nível Fundamental (5-8 anos)	Com marcas não-padrão (-EM, -U)		Sem marca (variante sem ditongo P3=P6)	
	Ocorrências	%	Ocorrências	%
Mulheres	36/52	69, 2	14/25	56
Homens	16/52	30, 8	11/25	44

[12] Linguagem popular ou à "vilão" [viˈʎẽw̃] tem sido objeto de interesse por parte de estudiosos madeirenses, como dão conta as publicações, entre outras, *Linguagem Popular da Madeira* de Eduardo Antonino Pestana (revistas Língua Portuguesa e Revista Lusitana, entre 1914 e 1940, e posteriormente editado em 1970, num volume intitulado Ilha da Madeira II - Estudos Madeirenses.), o *Vocabulário Popular do Arquipélago da Madeira*, do Padre Fernando Augusto da Silva (1950), ou ainda *Falares da Ilha - Dicionário da Linguagem Popular Madeirense*, de Abel Caldeira, na sua 2ª edição (1993).

Esses dados empíricos, tal como defendido por Lucchesi (2012) quando confrontado com o papel da variável *sexo* em variedades urbanas do PB, devem ser entendidos em função do papel social atribuído aos gêneros nas configurações socio-históricas específicas a cada comunidade. No caso do Brasil, Lucchesi (2012, p. 709) adianta estar em curso um nivelamento linguístico, através da influência exercida pelos grandes centros urbanos sobre as outras regiões, e que nesse contexto:

> os processos de variação socialmente marcados (tais como os da concordância nominal e verbal) apontam para processos de "cima pra baixo" e de fora para dentro dessas comunidades rurais e da periferia das grandes cidades brasileiras. Dessa forma, os homens, que em média, estão melhor inseridos no mercado de trabalho e circulam mais para fora da comunidade, tendem a assimilar primeiro as formas do padrão urbano, abandonando mais rapidamente o uso das formas desviantes da linguagem rural e popular.

A fraca implementação da variável concordância verbal em informantes pouco escolarizados, fenômeno não estatisticamente atestado nas variedades continentais do PE, mas amplamente estudado em variedades do PB, tem paralelo com um outro estudo realizado por Vianna (2011) sobre a implementação de *a gente* em variedades do Português. Nesse estudo, o fator *localidade* teve relevância estatística, distinguindo-se claramente a variedade insular das duas variedades continentais do PE (Cacém e Oeiras). O fator *escolaridade* atuou de forma determinante na amostra de Funchal, que apresenta 52% de ocorrências de *a gente* em informantes poucos escolarizados, contra apenas 32% no Cacém e 13% em Oeiras. A variável *sexo*, tal como verificado na implementação da concordância verbal de terceira pessoa, também condiciona a escolha da variante *a gente*, contrastando o comportamento dos informantes de Funchal com o das outras variedades urbanas do PE. O contraste concretiza-se na preferência feminina pela forma *a gente* (51%), a porcentagem mais elevada (Cacém, 30%; Oeiras 8%), enquanto os homens de Funchal preferem a forma prestigiada *nós* (91%), valores similares aos obtidos nas outras variedades (Cacém, 86%; Oeiras, 91%).

Andrade (1990), na sua análise do fenômeno da palatalização do /l/, variante não-padrão identificada como sendo um dos traços mais característicos da variedade insular, cuja opcionalidade está condicionada pela presença de [i] ou [j], oral e nasal, imediatamente anterior, quer pertençam à

palavra ou não [*quilo* > [kiʎu], *vi logo* [viʎɔgu], também verificou a ação de fatores extralinguísticos. Nesse caso também, e tendo em consideração unicamente o sexo, verifica-se que "globalmente, as mulheres palatizam mais do que os homens em todos os estilos considerados" (ANDRADE, 1990, p. 28).

Cabe assinalar que, no âmbito deste estudo, a existência de padrões flexionais variantes de P6 poderá resultar da atuação de variáveis extralinguísticas sociais (*escolaridade* e *sexo*), relacionadas ou não com aspectos históricos de construção dialetal, que se inicia no século XV, com o povoamento da Ilha da Madeira, no "ciclo de expansão da língua" (CASTRO, 2006), em descontinuidade com os processos de formação de variedades padrão e dialetais em Portugal Continental.

Paralelamente à formação da variedade padrão do PE, com base numa distanciação "não só do castelhano mas também dos dialetos setentrionais falados pela antiga nobreza portuguesa" (CARDEIRA, 2009, p. 111) e com o contributo da expansão portuguesa (através do contato linguístico em espaços não europeus mas também dentro do território português com a presença de escravos africanos, e cuja "presença na sociedade teria, forçosamente, algum impacto no favorecimento da simplificação e nivelamento linguísticos, fortalecendo a deriva do português meridional." (CARDEIRA, 2009, p. 112)), o Português falado pelos primeiros povoadores, a que se sucederam outras vagas de emigrantes oriundos de vários pontos de Portugal, entrou em contato[13], desde o início da sua implementação no espaço insular, com vários povos de grande diversidade étnica (europeus, atraídos pelas possibilidades de desenvolvimento mercantil, genoveses e florentinos, franceses, espanhóis, ingleses, mas também de escravos, guanches, mouros e negros, capturados para servirem de mão-de-obra) (VIEIRA, 1997; 2001).

Dados os limites da descrição sincrônica proposta neste estudo, o desenvolvimento dado à análise das variantes de concordância verbal na variedade insular do PE constitui uma tentativa de aprofundamento, apontada por Galves

[13] Sobre o povoamento da Ilha da Madeira e dados relativos à caracterização populacional nos séculos XV a XVIII, ver os estudos de Pinto; Rodrigues (1993) e Ferraz (1990). Em ambos, são referidas porcentagens aproximativas do número de escravos residentes na Ilha da Madeira durante esse período, sendo estes valores considerados significativos. Este contato também aconteceu em Portugal continental, sobretudo na cidade de Lisboa, capital do Reino, mas é provável que o impacto tenha sido maior num núcleo populacional de menores dimensões como o do Funchal e com uma estrutura social distinta (número reduzido de indivíduos nobres, denominados os "homens bons") (GONÇALVES, 1958).

(2012) como necessário, dos chamados dialetos europeus não-padrão. Por fim, no que se refere aos condicionamentos extralinguísticos, as variáveis escolaridade e sexo parecem constituir-se como marcadores diferenciadores da variedade insular relativamente ao PE padrão, explicáveis por condições específicas, de natureza histórico-política e social, de construção social da comunidade insular.

Atendendo, ainda, à proposta geral da presente obra, de abordar a interface morfossintaxe-fonética na descrição da concordância verbal de P6, procurou-se aprofundar, nesta seção, o comportamento dos fatores fonéticos que supostamente afetariam a marcação de pluralidade em PE. O exame detalhado das construções das três amostras europeias confirma que "a saliência fônica em si não parece desempenhar nenhum papel no português dialetal" (GALVES, 2012, p. 134), não se observando, por exemplo, ocorrências de cancelamento crescente das marcas plurais, como a evitar exemplos do tipo *eles vai, eles cantou* – resultado que parece colaborar para a interpretação da natureza da regra de concordância nas variedades do Português.

2.2. PADRÕES DE CONCORDÂNCIA - PB E PST: RESTRIÇÕES PARA A VARIAÇÃO

Nas comunidades situadas fora dos atuais domínios territoriais de Portugal, verifica-se, em maior ou menor dimensão, uma regra variável de concordância de P6.

Na variedade brasileira, os índices encontrados – 78,2% em NIG; 88,1% em COP – são compatíveis com as tendências verificadas em outros estudos da fala de comunidades urbanas, que costumam assumir comportamento linguístico bastante distinto do atestado em áreas rurais. O registro de altos índices de concordância em amostras brasileiras é característico especificamente de variedades urbanas e fundamentalmente praticadas por indivíduos escolarizados, o que contrasta fortemente com a situação de outras comunidades que se localizam no eixo rural ou rurbano (cf. BORTONI-RICARDO, 2005).

A estratificação social brasileira, flagrante no que se refere aos graus de escolarização e às classes sociais, de um lado, e ao perfil dos indivíduos quanto aos traços rural e urbano, de outro, revela-se linguisticamente em um *continuum* de maior ou menor realização das variantes consideradas portadoras de prestígio, dentre as quais avultam as referentes às marcas de

número. Em comunidades de falantes pouco escolarizados e/ou mais rurais, os índices de marcação de plural costumam ser bastante diferenciados dos relativos aos grupos de falantes muito escolarizados e/ou mais urbanos.

Em São Tomé e Príncipe, as taxas especialmente expressivas de marcação de plural, de 92,1%, resultam, ao que tudo indica, da forte influência da norma europeia de concordância, que serve de modelo para a aprendizagem da língua portuguesa como L2.

Resguardadas as diferenças sócio-históricas e linguísticas que envolvem as variedades são-tomense e brasileira, é importante verificar que a regra variável de concordância parece obedecer, hoje, a restrições estruturais e sociais muito semelhantes. A seleção de variáveis relevantes segundo a análise estatística, conforme se pode observar no quadro a seguir, evidencia essa semelhança.

Quadro 2. Variáveis atuantes para a implementação da (não) marcação de número em formas verbais de P6 por amostra/variedade

Copacabana	Nova Iguaçu	São Tomé
Escolaridade	Paralelismo (discursivo e oracional)	Paralelismo (discursivo e oracional)
Paralelismo (discursivo e oracional)	Configuração do sujeito	Posição do sujeito
Saliência fônica	Saliência fônica	Escolaridade
Faixa etária	Escolaridade	Sexo
Sexo	Posição do sujeito	Traço semântico do sujeito
Posição do sujeito	Traço semântico do sujeito	
Significância: .00	Significância: .02	Significância: .01
Input: .97	Input: .90-.88	Input: .97

Embora em ordem diferente, diversas variáveis – paralelismo oracional e discursivo, posição do sujeito, escolaridade – atuam tanto nas amostras brasileiras quanto na africana. Em segundo lugar, verifica-se a atuação dos grupos de fatores traço semântico do sujeito e sexo do informante, que se mostraram relevantes em duas das três amostras. Por fim, as variedades brasileiras se particularizam quanto à extensão relativa à diversidade de contextos em que se dá a alternância singular-plural, motivo pelo qual apenas nelas aparecem determinadas variáveis estruturais, como a saliência fônica, tanto em COP quanto em NIG. Por motivos referentes à delimitação do trabalho, as variáveis que ocorreram exclusivamente em um dos pontos de inquérito, como faixa etária em COP e configuração do sujeito em NIG, não serão descritas no presente capítulo.

2.2.1. O COMPORTAMENTO DAS VARIÁVEIS SOCIAIS

a) Escolaridade

É bastante esclarecedor e sinalizador da realidade social o efeito da escolaridade nas variedades brasileiras e africana, diferentemente do que se pode observar particularmente no PE continental, em que ser ou não escolarizado não faz alterar a regra semicategórica de concordância.

Tabela 11. Atuação da variável *Escolaridade* para a implementação da marca verbal de P6 por variedade/amostra

Variedade / Amostra		Nível fundamental (5 a 8 anos)	Nível médio (9 a 11 anos)	Nível superior (12 a 15 anos)
P B	COP	299/410 72,9% .17	340/382 89,0% .41	590/603 97,8% .79
P B	NIG	333/457 72,9% .32	243/362 67,1% .35	491/546 89,9% .68
P S T	ST	179/215 83,3% .14	220/236 93,2% .49	280/286 97,9% .81

Em linhas gerais, pode-se dizer que o efeito da escolarização, seguramente verificado nos pesos relativos, favorece nitidamente a concordância de forma crescente nas três variedades, embora com perfil levemente diferenciado em NIG. Enquanto o efeito crescente da escolaridade pode ser nitidamente observado nos índices de marcação plural em COP, região de alto prestígio socioeconômico, o que efetivamente demarca a mudança de comportamento

em Nova Iguaçu, região de baixo prestígio socioeconômico, é o acesso ao nível superior. De todo modo, a influência do acesso à escolarização tem sido fortemente atestada em diversas investigações, seja para sustentar a polarização sociolinguística brasileira (LUCCHESI, BAXTER, SILVA, 2009), seja para refletir sobre estabilidade e/ou mudança quanto à produtividade da concordância (NARO; SCHERRE, 2003). A esse respeito, os autores têm demonstrado que a urbanização e o acesso cada vez maior à escola acarretam o aumento, em diversas classes sociais, do uso das marcas de concordância.

Em São Tomé, em que a variável escolaridade atua de forma quase idêntica ao que ocorre em Copacabana, cabe salientar que o acesso à instrução implica, na realidade, maior uso da própria Língua Portuguesa, frente às línguas maternas adquiridas pelos informantes. Por essa razão, controlou-se, em particular, o maior ou menor contato com o Português, por meio do depoimento dos próprios informantes. Os resultados a seguir, que se referem à totalidade de dados produzidos por informantes são-tomenses, demonstram que a frequência de uso da língua crioula está intimamente correlacionada à realização da marca de concordância: quanto maior essa frequência, menor o uso das marcas plurais.

A título de esclarecimento, cabe informar que todos os indivíduos que atuaram na composição da amostra são-tomense se declararam falantes de Português como L1. Tendo em vista, entretanto, a pluralidade linguística na sociedade em questão, em que coexistem três crioulos, além de outras línguas minoritárias (cf. HAGEMEIER, 2009), e em que o uso da língua crioula é desprestigiado, o presente estudo controlou a possível influência do contato com o Português de acordo com a proposta de Brandão (2011), segundo a qual há indivíduos que declaram (a) só usar o português e dominar ou não dominar um crioulo (zero ou baixa frequência); (b) usar o português e um ou mais crioulos com menor frequência; ou (c) com maior frequência – no mercado, em momentos de lazer com amigos (alta frequência).

Tabela 12. Atuação da variável *Frequência de uso de um crioulo* para a implementação da marca verbal de P6 no PST

Frequência de uso de um crioulo	Oco	Perc.	P.R.
Zero /Baixa	465/503	92.4%	.57
Média	572/634	90.2%	.48
Alta	71/89	79.8%	.35

Tendo em vista, ainda, a íntima relação entre acesso a escolaridade e maior uso da Língua Portuguesa na sociedade de São Tomé, realizou-se o cruzamento dos dados referente às variáveis *frequência de uso de um crioulo* e *escolaridade*.

Tabela 13. *Frequência de uso de um crioulo* versus *Nível de escolaridade* quanto à realização da marca verbal de P6 no PST

Frequência de uso de um crioulo	Nível de escolaridade					
	Fundamental (5 a 8 anos)		Médio (9 a 11 anos)		Superior (12 a 15 anos)	
Zero/Baixa	122/139	88%	740/222	92%	138/142	97%
Média	181/217	83%	237/260	91%	154/157	98%
Alta	71/89	80%	----	-----	-----	----
Totais	374/445 - 84%		442/482 - 92,7%		292/299 - 97,7%	

Os dados revelam que nenhum dos informantes dos níveis médio e superior de escolaridade declarou usar um crioulo com alta frequência, o que confirma ser a instrução uma espécie de passaporte para o domínio da Língua Portuguesa na sociedade são-tomense. Entre os informantes de nível fundamental, confirma-se que, de forma escalar, quanto maior a frequência de uso de um crioulo, menores são os índices de marcação verbal de P6.

Salienta-se, aqui, que a correlação entre as referidas variáveis, na qualidade de evidência empírica, favorece a hipótese de que o maior contato linguístico se correlacione à menor realização da marca, hipótese empregada nos estudos que defendem ser a não concordância brasileira, nos níveis e contextos existentes, oriunda da forte situação de contato linguístico entre negros, índios e brancos.

b) Sexo do informante

Conforme registrado anteriormente, apenas para a amostra referente a NIG, não se verificou comportamento estatisticamente relevante, entre os índices de concordância registrados para homens e mulheres, cuja diferença é de menos de 3%. Nas localidades de Copacabana e São Tomé, assim como ocorreu com o fenômeno da concordância nominal, a diferença do comportamento entre homens e mulheres, conforme se pode verificar na tabela a seguir, mostrou-se estatisticamente relevante.

Tabela 14. Atuação da variável *Sexo* para a implementação da marca verbal de P6 por variedade/amostra

Variedade / Amostra		Masculino	Feminino
PB	COP	761/882 86,3% .43	468/513 91,2% .63
	NIG	483/607 79,6%	584/758 77,0%
PST	ST	380/424 89,6% .37	299/313 95,5% .68

Em COP e ST, atestou-se que as mulheres (.63; .68) tendem à maior realização das marcas de número do que os homens (.43; .37). Essa tendência confirma o postulado laboviano de que, em muitas sociedades, as mulheres, embora costumem ser as mais inovadoras em processos de mudança linguística, não o são quando se trata de uma variante de pouco prestígio social.

O fato de essa tendência não ter sido atestada em Nova Iguaçu, em que homens e mulheres têm comportamento aproximado, nem em Funchal, em que se verificou a tendência contrária, pode sinalizar a diferença de gênero em cada organização social. Ao que tudo indica, embora a hipótese deva ser validada com investigação de natureza sociológica específica, quanto maior a inserção da mulher na vida social, econômica, profissional (como parece ser o caso de Copacabana e São Tomé), maior a tendência a se efetivar o uso das marcas de plural, forte sinalizador de detenção da cultura letrada.

2.2.2. O COMPORTAMENTO DAS VARIÁVEIS LINGUÍSTICAS

a) Paralelismo discursivo (de verbo para verbo)

O efeito do princípio do paralelismo (cf. SCHERRE, 1988) no âmbito da concordância verbal – que postula que "marcas levam a marcas e zeros levam a zeros", seja do sujeito para o verbo (paralelismo oracional), seja de um verbo para outro verbo (paralelismo discursivo) – já foi vastamente atestado em amostras do PB. Os resultados expostos na tabela a seguir não só confirmam tal efeito em dados brasileiros, mas também o atestam na amostra são-tomense.

Tabela 15. Atuação da variável *Paralelismo discursivo* para a implementação da marca verbal de P6 por variedade/amostra

Variedade / Amostra		verbo antec. de v. com marca	verbo antec. de v. sem marca	v. primeiro de série ou isolado
PB	COP	359/359 100%	6/74 8,1%	864/962 89,8%
	NIG	269/277 97,1% .92	4/121 3,3% .01	794/967 82,1% .52
PST	ST	171/171 100%	15/15 0%	508/551 92,2%

O efeito paralelístico da presença / ausência de marcas de plural nos verbos em série é nítido no PB e no PST, o que é evidenciado seja pelo efeito categórico (COP e ST), seja pelos discrepantes pesos relativos (.92 *versus* .01, em NIG).

Ao que tudo indica, em qualquer amostra em que minimamente se verifique a alternância entre as formas singular e plural, pode-se observar o efeito do paralelismo, por ser esse grupo de fatores, por hipótese, de natureza mais cognitivo-processual, referente ao processamento e à memória da informação, do que estritamente linguística. Nesse sentido, embora não tenha sido possível observar sua atuação no PE continental, por efetiva falta de variação, os dados de Funchal, ainda que em número reduzido, sugeriram tal efeito.

b) Paralelismo oracional (do sujeito para o verbo)

Diferentemente do verificado nas amostras europeias – em que nem sequer foi possível observar o efeito das marcas do sujeito sobre as do verbo, por ser a concordância nominal categórica (cf. BRANDÃO, 2013) –, o paralelismo oracional também demonstrou forte atuação nas amostras brasileiras e são-tomense.

Tabela 16. Atuação da variável *Paralelismo oracional* para a implementação da marca verbal de P6 por variedade/amostra

Variedade / Amostra		verbo antec. de suj. com marca	verbo antec. de suj. sem marca	v. antec. de sujeito numeral
P B	COP	918/991 92,6% .54	13/33 39,4% .14	5/7 71,4% ---
P B	NIG	807/954 84,6% .54	29/65 44,6% .14	7/11 63,6% ---
P S T	ST	462/490 94,3% .52	8/15 53,3% .17	1/1 100% ----

Conforme indicam os pesos relativos, idênticos ou quase idênticos, o efeito paralelo da presença/ausência das marcas de plural no sujeito é evidente tanto nas amostras do PB quanto na de São Tomé. De fato, sujeitos não marcados acarretam desfavorecimento (.14; .17) da marcação plural no verbo, comparados aos marcados (.54; .52).

c) Posição do sujeito

Conforme os resultados na tabela a seguir, é evidente o efeito da posposição do sujeito no desfavorecimento da marcação plural (.16 em COP; .17 em NIG; e .08 em ST), contrariamente ao que ocorre nos contextos com sujeito anteposto, com índices superiores ao ponto neutro, de .50.

Tabela 17. Atuação da variável *Posição do sujeito em relação ao verbo* para a implementação da marca verbal de P6 por variedade/amostra

Variedade / Amostra		Sujeito anteposto	Sujeito posposto
P B	COP	916/1007 91,0% .52	37/44 84,1% .16
P B	NIG	830/1003 82,8% .52	18/39 46,2% .17
P S T	ST	462/492 93,9% .66	29/39 74,4% .08

A relação dessa variável com a transitividade é sempre aventada como hipótese: ao que parece, há uma correlação lógica entre sujeitos pospostos e inacusatividade, em primeiro lugar, e intransitividade, em segundo. Estatisticamente, entretanto, não só neste estudo, mas também em trabalhos de outros pesquisadores (SCHERRE; NARO, 2007) fica comprovado que, em amostras brasileiras, o efeito de posição costuma ser superior ao de transitividade:

> (...), independentemente do tipo de verbo, qualquer argumento ou sintagma à direita do verbo tende, relativamente, a diminuir as marcas de concordância explícita. (SCHERRE; NARO, 2007, p. 312)

Conforme revelam os dados da presente pesquisa, mesmo no PE continental (CARDOSO; CARRILHO; PEREIRA, 2011), mas principalmente nas entrevistas em Funchal, esse efeito é verificado. Na realidade, conforme demonstra Corbett (2000), a posição do sujeito (além de seu traço semântico de animacidade) altera padrões de concordância em línguas de perfis e genealogias muito diferentes. Esse dado faz supor que não se trate de efeito efetivamente variável nesse caso, mas de comportamento estrutural gramaticalmente diferenciado.

Por essa razão, a presente investigação optou por observar os dados ora com a presença dos sujeitos pospostos, ora sem eles. Em linhas gerais, deve-se advertir que, embora o conjunto das demais variáveis relevantes

atue na presença ou na ausência de tais estruturas, a exclusão dos dados de posposição de sujeito foi fundamental para atestar a particularidade das amostras europeias, no sentido de que a aparente variação em Funchal, por exemplo, resulta, em boa parte, da presença de sujeitos pospostos, conforme já se demonstrou em seção anterior.

d) Traço semântico do sujeito

Os resultados, a seguir, demonstram a influência do traço semântico do sujeito referente ao caráter mais/menos animado do referente nas amostras de Nova Iguaçu e São Tomé.

Tabela 18. Atuação da variável *Traço semântico do sujeito* para a implementação da marca verbal de P6 por variedade/amostra

Variedade / Amostra		Sujeito animado	Sujeito inanimado
P B	COP	1035/1175 88,1%	194/220 88,2%
	NIG	948/1187 79,9% .53	119/178 66,9% .36
P S T	ST	575/612 94,0% .56	104/125 83,2% .26

Confirma-se, com base nos pesos relativos quase idênticos, a conhecida influência do traço semântico do sujeito: sujeitos inanimados desfavorecem a marca de número plural (.36 em NIG e .26 em ST), enquanto sujeitos animados a favorecem (.53 em NIG e .56 em ST).

De natureza semântico-discursiva, cabe destacar a atuação dessa variável, que supõe haver uma saliência no processamento da informação, segundo a qual seria sublinhado, por meio da concordância, o traço dinâmico (de controle) nas relações inauguradas pelos predicadores verbais. Não por acaso esse efeito também pôde ser minimamente verificado nas amostras europeias e já foi atestado em diferentes línguas do mundo.

e) Saliência fônica

O princípio da saliência fônica (LEMLE; NARO, 1977) – segundo o qual quanto maior a diferenciação fônica entre as formas singular e plural, maior a tendência à concordância – não alcança a "universalidade" aparentemente verificável na atuação dos grupos de fatores linguísticos anteriormente tratados. Os resultados da presente investigação revelam que o efeito da saliência fônica ocorreu exclusivamente nas variedades brasileiras.

Tabela 19. Atuação da variável *Saliência fônica* para a implementação da marca verbal de P6 por variedade/amostra

Variedade / Amostra		Grau 1	Grau 2	Grau 3	Grau 4
P B	COP	669/795 84,2% .35	160/173 92,5% .51	180/192 93,8% .79	220/235 93,6% .70
	NIG	592/812 72,9% .29	193/245 78,8% .55	144/156 92,3% .78	138/152 90,8% .74

Excetuando-se os graus 3 e 4, em que não se verificaram pesos relativos crescentes de marcação plural, comprovou-se o princípio geral segundo o qual a maior diferenciação fônica entre as formas singular e plural dos verbos acarreta, de forma crescente, maior tendência à concordância (em COP, .35 > .51 > .79/.70; em NIG, .29, .55, .78/.74).

Diversos trabalhos atestaram a relevância dessa variável em amostras brasileiras. É curioso observar, entretanto, como bem pontua o trabalho de Scherre, Naro (2006), que esse grupo de fatores não costuma se revelar estatisticamente relevante quando se trata de dados produzidos exclusivamente por falantes escolarizados. Ao que parece, assim como ocorre nas variedades europeias e são-tomenses, a preferência massiva pela concordância faz com que os casos de não marcação plural estejam sobremaneira concentrados nos níveis iniciais de diferenciação fônica que o efeito crescente dos graus de saliência nem sequer possa ser observado. Para maior aprofundamento dessa variável no que se refere ao condicionamento fonético, vale conferir o Capítulo 4 da presente obra.

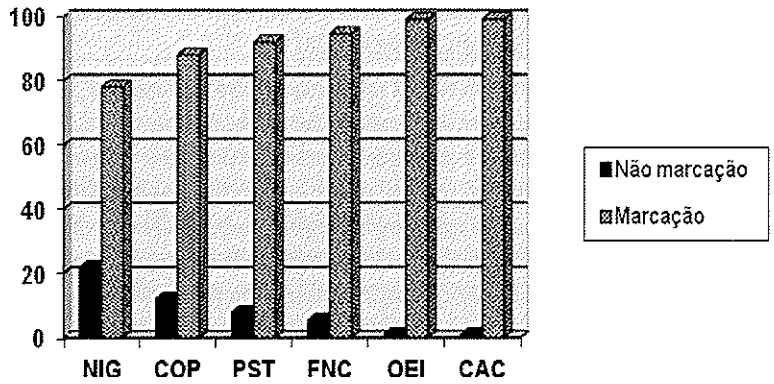
3. SISTEMATIZAÇÃO E ABORDAGEM CONTRASTIVA DOS RESULTADOS DA CONCORDÂNCIA VERBAL DE P6 NAS VARIEDADES DO PORTUGUÊS EM ANÁLISE

Com base nos resultados ora apresentados, associados, quando possível, aos de outros estudos sobre o tema, propõe-se, nesta seção, uma abordagem contrastiva entre os padrões de concordância de P6 verificados nas variedades do Português.

3.1. DA REALIZAÇÃO DA CONCORDÂNCIA NAS VARIEDADES DO PORTUGUÊS: ASPECTOS QUANTITATIVOS

Valendo-se exclusivamente das frequências brutas de marcação de P6 verificadas no presente estudo, fica patente que, de todas as variedades urbanas, todas elas partilhando o prestígio promovido pelos meios letrados e escolarizados, os padrões europeus se diferenciam dos demais. O gráfico a seguir facilita visualizar a diferença.

Figura 1. Índices de marcação e não marcação de plural em verbos de P6 nas amostras brasileiras (NIG e COP), africana (PST) e portuguesas (FNC, OEI, CAC)



A imagem, por si só, permite detectar o *continuum* que vai do maior afastamento do polo da concordância, representado pelas variedades brasileiras, até o menor afastamento desse polo, representado pelas variedades continentais do PE, de índices nitidamente semicategóricos.

Em porção intermediária, ficam as variedades de São Tomé e Funchal, aquela com expressão das formas alternantes em uma diversidade de contextos bem mais extensa do que esta.

À representatividade desses dados para a caracterização das referidas variedades poderiam se opor resultados contrários advindos da descrição de outras amostras. Por essa razão, cabe, ainda que brevemente, comparar os índices gerais obtidos aos de outras investigações.

Embora seja muito relevante a existência de padrões de concordância no PST semelhantes aos do PB, não só no âmbito da concordância verbal mas também no da nominal (Cf. BRANDÃO; VIEIRA, 2012a), especialmente por servir como evidência da perda de marcas flexionais típicas da situação de intenso contato linguístico, o pouco conhecimento que se tem sobre a expressão da terceira pessoa nas variedades africanas do Português não enseja maior debate nesta seção.

No PB, a realidade variável da concordância já foi amplamente descrita em amostras representativas de diferentes comunidades de fala, em termos não só regionais, mas também sociais. Como não caberia nos limites do presente texto uma apresentação detalhada desses estudos, sistematizam-se, no quadro adiante, os índices gerais obtidos em investigações que demonstrem o amplo espectro da variação configurada em um contínuo do mais urbano ao mais rural no Estado do Rio de Janeiro, o que se correlaciona, ainda, ao grau de instrução do informante.

Naro (1981) e Vieira (1995) ocuparam-se da fala de analfabetos; enquanto o primeiro trabalho se vale da fala da região metropolitana, com a amostra MOBRAL, na década de 1970, o último investiga a fala de pescadores da Região Norte fluminense, com o *Corpus* APERJ, na década de 1980. Além da presente investigação, com dados contemporâneos recolhidos entre 2007 e 2011, Scherre; Naro (2006) observa a fala de cariocas com escolaridade fundamental e média a partir de entrevistas feitas no período aproximado a 1980, e ainda contempla dados extraídos de entrevistas feitas no período aproximado ao fim do século XX, e Graciosa (1991) trata da fala de indivíduos com curso superior, com base no *Corpus* NURC-RJ.

Tabela 20. Frequência de marcação de plural em diferentes estudos sobre a concordância verbal de 3ª p. plural no PB (Estado do Rio de Janeiro)

Variedade			Localidade	Nível de escolaridade dos falantes	Marcação de plural
BRASIL	URBANA		Amostra Mobral - 1970 (NARO, 1981)	Todos os dados Analfabetos	48%
			PEUL - 1980 (SCHERRE; NARO 2006)	Todos os dados	73%
				1-4 anos	63%
				5-8 anos	78%
				9-11 anos	81%
			Nurc/RJ – 1970/80 (GRACIOSA, 1991)	Superior	94%
			PEUL - 2000 (SCHERRE; NARO 2006)	Todos os dados	84%
				1-4 anos	75%
				5-8 anos	85%
				9-11 anos	93%
			Copacabana - 2010 (VIEIRA; BAZENGA, 2013)	Todos os dados	88,1%
				5-8 anos	72,9%
				9-11 anos	89,0%
				12-15 anos	97,8%
			Nova Iguaçu - 2010 (VIEIRA; BAZENGA, 2013)	Todos os dados	78,2%
				5-8 anos	72,9%
				9-11 anos	67,1%
				12-15 anos	89,9%
		NÃO URBANA	*Corpus* APERJ - 1980 (VIEIRA, 1995)	Todos os dados (0-4 anos)	38%

Os dados de P6 coletados no Estado do Rio de Janeiro, na segunda metade do século XX ou no início deste, revelam índices crescentes de concordância num *continuum* que vai do mais rural ao mais urbano, e do menos escolarizado ao mais escolarizado. Esse *continuum* é proposto por Lucchesi, Baxter; Silva (2009, p. 348), que, valendo-se de resultados de outros estudos brasileiros e incluindo os relativos à fala da comunidade afrobrasileira de Helvécia/Bahia, consideravelmente isolada, interpretam a distribuição dos dados como reveladora da polarização sociolinguística brasileira. O esquema a seguir, que inclui as taxas gerais de concordância expostas na Tabela 20 e ainda os resultados de Helvécia, também representa tal *continuum*.

Figura 2. *Continuum* de marcação de pluralidade em verbos de P6 consoante os traços rural / urbano e os graus de escolaridade em variedades brasileiras com base na proposta de Lucchesi; Baxter; Silva (2009, p. 348)

[+ marcas]				[- marcas]
RJ	RJ	RJ	RJ	Helvécia/BA
Urbano	Urbano	Urbano	Rural	Rural
Altamente escolarizado	Moderadamente escolarizado	Analfabeto	Analfabeto	Comunidades afrobrasileiras
94 / 97,8%	73 / 89%	48%	38%	16%

No que se refere às poucas descrições de amostras do PE, destacam-se aqui os seguintes trabalhos, os únicos a que se teve acesso[14]: Varejão (2006) e Gandra (2009), ambos com base em entrevistas do *Corpus* CORDIAL-SIN, que contempla as chamadas variedades dialetais (populares) do Português falado em toda a extensão de Portugal e ilhas; Monguilhott (2009), que considera falantes com escolaridade fundamental ou média nascidos na Grande Lisboa; Rubio (2012), que descreve a fala de informantes também de diversas regiões de Portugal, do *Corpus* de Referências do Português Contemporâneo (CRPC), os quais foram distribuídos segundo quatro graus de escolaridade; e Monte (2012), que considera informantes analfabetos, com escolaridade secundária (do tipo Educação de Jovens e Adultos – EJA) e superior, da região de Évora. A tabela abaixo sintetiza os percentuais de marcação de P6 em cada estudo.

[14] Foram selecionados apenas trabalhos de outros pesquisadores, tendo em vista que os trabalhos anteriores feitos pelas autoras do presente trabalho contavam com parte da amostra básica do *Corpus* Concordância e foram reconfigurados no presente artigo.

Tabela 21. Frequência de marcação de plural em diferentes estudos sobre a concordância verbal de P6 no PE

Variedade		Localidade	Nível de escolaridade dos falantes	Marcação de plural
PORTUGAL	URBANA	Grande Lisboa[1] (MONGUILHOTT, 2009)	**Todos os dados**	**91%**
			Fundamental	89-90%
			Superior	92-94%
		Évora (MONTE, 2012)	**Todos os dados**	**93,1%**
			Não alfabetizados	91,2%
			Ensino básico (EJA)	93,5%
			Ensino secundário	94,4%
		Oeiras/Lisboa; Cacém (VIEIRA; BAZENGA, 2013)	**Todos os dados**	**99,2%**
			Ensino fundamental	98,9%
			Ensino médio	99,6%
			Ensino superior	99,1%
		Funchal (VIEIRA; BAZENGA, 2013)	**Todos os dados**	**94,7%**
			Ensino fundamental	89,8%
			Ensino médio	95,9%
			Ensino superior	96,8%
	NÃO URBANA	Diversos pontos – Cordial-Sin (VAREJÃO, 2006; GANDRA, 2009[2])	**Todos os dados** (0-4 anos)	**91%; 96,5%**
		Diversos pontos - CRPC (RUBIO, 2012)	**Todos os dados**	**93,9%**
			1-4 anos	91,9%
			5-8 anos	93,9%
			9-11 anos	93,7%
			12 ou mais	96,2%

Como se pode observar no quadro acima, constata-se a alta realização da marca de plural em todos os estudos. Independentemente da região e do perfil mais urbano ou mais rural da comunidade, todas as amostras apresentam mais de 90% de realização da marca. Tendo em vista o limiar utilizado na presente investigação para diferenciar uma regra variável (até 94% aproximadamente) de uma semicategórica (a partir de 95%), conforme Labov (2003), cabe discutir a existência de registros de concordância na faixa de 91 a 94%, o que superficialmente diferenciaria os resultados deste trabalho e os dos referidos estudos.

No que se refere aos índices relativos aos níveis de escolaridade, a expressiva proximidade entre eles impede propor com segurança o efeito dessa variável. Excetuando-se o observado em Funchal, cujas características particulares já foram discutidas anteriormente, os percentuais não distam nem cinco pontos. Considerando a totalidade dos estudos, confirma-se a preferência pela concordância independentemente do grau de instrução do indivíduo.

Observando especificamente as orientações metodológicas de cada investigação, fica patente que as diferenças quantitativas resultam dos procedimentos adotados quanto à coleta de dados. Confirma o teor dessa afirmativa o fato de que trabalhos que se valem do mesmo *corpus* (embora com localidades nem sempre equivalentes), como é o caso de Gandra (2009) e de Varejão (2006), alcançaram resultados diferenciados: 91 e 96,5%, respectivamente[15]. Sem dúvida, advogar em defesa da inclusão e da exclusão de tais dados implica assumir posições referentes à interpretação das estruturas que envolvem a terceira pessoa. Por essa razão, apresenta-se, na próxima seção, uma abordagem qualitativa dos contextos que foram verificados nos estudos de cada variedade.

3.2. COMPORTAMENTOS DAS VARIÁVEIS E A DEFINIÇÃO DAS VARIEDADES: ASPECTOS QUALITATIVOS

Em linhas gerais, enquanto nas amostras brasileiras ocorrem a realização e o cancelamento da marca plural em diversificados contextos estruturais, e em quantidades semelhantes em função sobretudo da escolaridade e do perfil rural/urbano, conforme já se demonstrou, é preciso avaliar, de modo detalhado, os resultados referentes ao PE, que se mostraram em alguns aspectos divergentes. Com base nessa avaliação, é possível, então, pôr em debate a proposta, veiculada em estudos como os de Naro; Scherre (2007), de que as diferenças entre as variedades brasileira e europeia do Português são mais quantitativas do que qualitativas.

Para definir, então, a extensão qualitativa dos contextos variáveis em amostras europeias, três procedimentos foram adotados: (i) o levantamento dos tipos estruturais contemplados na coleta dos dados; (ii) a observação das ocorrências de verbos não marcados; e (iii) a verificação do perfil das variáveis que se mostraram estatisticamente relevantes em trabalhos de orientação laboviana.

[15] Da mesma forma, Bazenga (2012), que também se valeu de dados do *Corpus* Concordância, havia registrado anteriormente mais de 10% de não concordância na variedade madeirense, índice que se relaciona fundamentalmente ao fato de ter incluído as ocorrências do tipo *tem/têm* e construções com verbo *ser* antecedido de SN topicalizado.

3.2.1. A COLETA DE DADOS PARA O ESTUDO DA CONCORDÂNCIA DE P6

Indicam-se, no quadro a seguir, os contextos que, segundo declarações dos autores, foram excluídos da coleta de dados dos trabalhos com amostras europeias.

Quadro 3. Listagem dos critérios de exclusão de dados nos estudos de amostras europeias

Contexto estrutural excluído da coleta de dados	Estudo sociolinguístico
Verbos de P6 em contextos de indeterminação do sujeito	Monguilhott (2009) Monte (2012) Rubio (2012)
Estrutura tradicionalmente chamada de passiva sintética	Monguilhott (2009) Varejão (2006)
Concordância semântica com nome singular (*o pessoal*)	Monguilhott (2009) Monte (2012)
Formas verbais homófonas do tipo *tem/têm*	Monte (2012) Varejão (2006) Rubio (2012)
Respostas em que se repete a forma verbal da pergunta feita pelo documentador	Monte (2012) Varejão (2006)
Orações com verbo 'ter' com valor existencial	Monte (2012)
Sujeito representado pelo pronome indefinido 'tudo' remetendo a um SN de 3ª pessoa do plural (os filhos e o marido é tudo Benfica)	Monte (2012)
Sujeito constituído por expressão partitiva	Monte (2012)
Verbos no infinitivo pessoal	Monte (2012)
Sujeito composto e posposto ao verbo com o núcleo mais próximo no singular (*morreu a senhora e o patrão*)	Monte (2012) Varejão (2006)
Sujeito do tipo pronomes interrogativos	Monte (2012)
Frases em que ocorre uma possível concordância com um pronome expletivo, do tipo *ele falta-me umas peças ...*	Varejão (2006)

Chama a atenção, em primeiro lugar, o diferenciado grau de detalhamento dos critérios de exclusão de dados empregados em cada investigação. É sem dúvida o trabalho de Monte (2012) o que mais explicita tais critérios. A rigor, se for tomado por pressuposto que os contextos não mencionados foram considerados na coleta de dados, os índices globais não poderiam ser seguramente comparados. Ocorre que, pela observação dos exemplos dados nos referidos estudos, pode-se supor que, embora alguns contextos

de exclusão não tenham sido mencionados, muito provavelmente os demais trabalhos também não os tenham considerado. Entretanto, há de fato estruturas interpretadas como variáveis em um estudo, e não variáveis em outro.

Rubio (2012), por exemplo, diferentemente de Monte (2012), computa normalmente ocorrências como a seguinte, em que o verbo também se relaciona com a referência genérica veiculada por "tudo" no contexto posterior à forma verbal:

> *(5) é. e depois há, há isto que, que parece-me, parece-me que tem importância, é que, dantes a farmácia era a farmácia oficina, on[de], onde ha(...), havia... como sabem...agora os medicamentos é quase tudo especializado* [CRPC-1082-20]

Em linhas gerais, o que está na base da exclusão dos dados é a impossibilidade de determinar a referência do sujeito como efetivamente plural, o que acaba por gerar uma construção ambígua por sua própria natureza, podendo o referente ser um singular ou um plural (numa espécie de concordância com o sentido), ou ainda indeterminado/genérico (construções de verbo na terceira pessoa do plural indeterminado; casos de infinitivo)[16].

Comparando os referidos trabalhos com o presente, este – embora tenha inicialmente contemplado a voz passiva sintética – se particulariza por estender o critério de exclusão de dados, ainda, aos seguintes casos: aparentes sujeitos de terceira pessoa plural (expresso ou nulo) que remetem a constituintes topicalizados; construções com verbo copulativo, que podem ser interpretadas como apresentacionais (*os caras é isso aí...*); construções cuja determinação da forma verbal como singular ou plural é de definição duvidosa por motivos fonético-fonológicos. Ao que tudo indica, está na interpretação de contextos como ambíguos, especialmente no que se refere aos dois primeiros tipos estruturais aqui referidos, a justificativa para os índices mais baixos de concordância nos demais estudos. Se todos os trabalhos citados eliminassem ou tratassem de forma particular, *a priori*, os contextos tomados aqui como ambíguos ou como de difícil determinação do estatuto

[16] Os diversos contextos de aparente variação de concordância são detalhadamente tratados em Martins (2012), que valida a exclusão de praticamente todos os casos do Quadro 3. Exceção feita aos dados de infinitivo (se pessoal ou impessoal), que não são mencionados pela referida autora, e os de pronomes interrogativos, que o próprio autor (MONTE, 2012) admite ter excluído fundamentalmente por serem em número muito reduzido.

singular ou plural do verbo, muito provavelmente seriam registradas nas referidas pesquisas taxas mais altas, o que confirmaria o estatuto semicategórico da concordância no PE.

Como se pôde observar dos critérios de coleta de dados, as quantidades absolutas apresentadas nos estudos já representam *a priori* a delimitação qualitativa do fenômeno da concordância, de modo que os elementos do par quantitativo – qualitativo são interdependentes. Nesse sentido, o estatuto supostamente variável do PE é definido pela interpretação estrutural das ocorrências em questão.

Para observar melhor a qualidade dos contextos encontrados em cada estudo, passa-se à verificação das ocorrências de não concordância usualmente encontradas em amostras europeias.

3.2.2. PERFIL DAS OCORRÊNCIAS DE VERBOS NÃO MARCADOS NO PE

No presente estudo, foi possível – ao listar todas as ocorrências sem a marca plural em Lisboa e em Cacém, e ao categorizar os dados sem a desinência explícita ou sem a desinência padrão em Funchal – controlar a qualidade das ocorrências do cancelamento da marca de número, à semelhança do que fez Vieira (2011). Do que se detalhou na descrição dos resultados, podem-se listar as construções mais comuns: (i) estruturas do tipo V SN compostas em sua maioria por verbos inacusativos; (ii) presença de sujeitos com traço semântico [- animado]; (iii) presença de verbo seguido de um fone (consoante nasal; vogal) que pode alterar a qualidade vocálica final do verbo e não se configurar, possivelmente, uma forma verbal idêntica à do singular.

Além desses contextos, mais produtivos, os exemplos encontrados neste ou noutros estudos registram a presença do relativo *que*, cujo favorecimento da forma singular, por inaugurar uma relação do verbo com seu referente mediada por uma forma não pluralizável, ou a presença de SNs complexos, influência que já foi seguramente detectada em outras investigações sobre o fenômeno.

Ademais, comparando os resultados da presente investigação aos de outras, verifica-se que abastecem os índices de cancelamento da marca de pluralidade, se não eliminadas pelo critério de coleta de dados, construções com verbo *ser* ou com outros verbos antecedidos de um SN que pode ser interpretado como tópico.

É importante observar que, mesmo coletando ocorrências dessa natureza, os estudiosos chegam a admitir que seus índices estejam intimamente relacionados a esses tipos estruturais específicos. Rubio (2012, p. 372) afirma que "para a concordância verbal de terceira pessoa do plural, a variação é semicategórica, visto restringir-se a apenas alguns contextos". Gandra (2009, p. 153) adverte que, nas ocorrências com verbo *ser*, ocorre constantemente "uma vacilação entre o sentido singular e plural do elemento a que se está fazendo referência". Monguilhot; Coelho (2012, p. 313) admitem que, em análise de dados em amostra lisboeta, o contexto sintático com verbo copulativo registra a concordância com o predicativo, e aventam a hipótese de que haveria na posição de sujeito um pronome nulo neutro (como um *isso*), como em "*os sonhos (isso) é assim umas bolinhas*". Monte (2012), reconhecendo a especificidade do fenômeno nas construções com verbo *ser*, chega a separar o *corpus* isolando-as das demais ocorrências verbais.

Já que são esses os tipos de construções que indicariam o cancelamento da marca de número no PE, é fundamental que se interprete o caráter variável ou não de tais estruturas. O presente trabalho reafirma a posição assumida anteriormente (Vieira, 2011) e faz coro às interpretações de que esses enunciados não evidenciam uma falta de concordância, ao menos nos mesmos moldes de construções não marcadas como as que ocorrem em frases brasileiras (e africanas) do tipo "*os menino comeu maçãs, mas nós não come essa fruta*". As particularidades dos contextos encontrados em amostras europeias localizam-se em três esferas: sintática, no caso do chamado sujeito posposto; semântica, nos casos que envolvem, em alguma medida, a ambiguidade na referência singular ou plural; e fonético-fonológica, nos enunciados em que as realizações das desinências número-pessoais põem em dúvida a equivalência ou a diferença entre as formas singular e plural.

Quanto aos casos de sujeitos supostamente pospostos, já se tem acesso à vasta literatura que questiona o estatuto do SN posposto ao verbo e advoga que ele tenha feições mais comparáveis às de um objeto direto, do que propriamente às de um sujeito. Monte (2009, p. 76-77) demonstra que diversos estudiosos defenderam essa tese. Trabalhos de cunho gerativista têm proposto ser essa construção resultante da associação de um verbo inacusativo com argumento interno, e não verdadeiramente um sujeito, que, nesse caso, seria um expletivo nulo. Seja nesses quadros teóricos ou em outros, é fundamental que se assuma o estatuto diferenciado dessas construções, que podem ocorrer em qualquer variedade do Português ou até de outras línguas, e nada esclarecem sobre as feições próprias da variedade em que se encontram.

Segundo Galves (2012), em artigo que propõe uma aprofundada reflexão sobre os fatos empíricos e as noções teóricas envolvidos no debate sobre as origens do PB, se resultante de "um processo de continuação do PE (deriva) ou se de um processo de ruptura (contato)" (p. 123), a análise das construções inacusativas se estende também àquelas com o verbo *ser*:

> A análise dos verbos inacusativos se estende facilmente ao verbo ser, com a diferença que esse verbo também toma orações como complementos. Todos os casos em que o verbo ser está no singular podem ser analisados como construções em que seu sujeito é um expletivo nulo, e a concordância se faz com esse expletivo. (GALVES, 2012, p. 131)

Além do estatuto sintático diferenciado das estruturas ora comentadas, há que se retomar os casos de ambiguidade na interpretação do sujeito como singular ou plural não só com *ser* mas também com outros verbos. Martins (2012), em artigo que detalha diversos casos em que um mesmo enunciado pode materializar estruturas sintáticas diferentes, configurando casos de ambiguidade estrutural, defende que

> Geram ambiguidade quanto à distinção singular/plural os sujeitos pronominais que retomam um constituinte topicalizado (podendo ser distintos os valores do traço de número do sintagma nominal topicalizado, por um lado, e do pronome sujeito, por outro). É o pronome de retoma do tópico, com a função de sujeito, que fixa a concordância sujeito-verbo, quer seja fonologicamente realizado, quer seja um pronome nulo. (MARTINS, 2012, p. 97)

Nesse sentido, a autora – em conformidade com o que sugere Mota; Vieira (2008) e propõe Martins (2012) – também defende que estruturas como "*Peras em vinho é a minha sobremesa favorita*" podem receber a interpretação equivalente a "*peras em vinho, [pronome nulo] é a minha sobremesa favorita*', sendo "peras em vinho" um tópico não sujeito. Galves (2012) também trata desse tipo de construção – como nas frases "*os nossos agasalhos [exp] é estes; duas canas [exp] dá para pescar*" – e, comparando-o com frases extraídas do francês, admite a análise segundo a qual o SN anteposto ao verbo seria o tópico da frase, e não o sujeito.

Contrariamente às aparências, não haveria, portanto, variação na concordância sujeito-verbo nesse e em outros casos; antes, ocorreria o que é postulado na tradição gramatical luso-brasileira no sentido de que a concordância sujeito-verbo pode ser estabelecida por vezes com a ideia, o sentido, como na chamada concordância ideológica.

Por fim, quanto aos enunciados em que as realizações das desinências número-pessoais em termos fonético-fonológicos põem em dúvida a equivalência ou a diferença entre as formas singular e plural, o reconhecimento das influências do contexto seguinte ao verbo (especialmente as consoantes nasais e as vogais) demonstradas neste estudo, bem como a descrição detalhada das pronúncias das desinências número-pessoais no PE (cf. MOTA; VIEIRA, 2008; MOTA, 2013; BAZENGA, Capítulo 3 deste livro) deixam clara a necessidade de descrever com mais rigor os crescentes graus de diferenciação fônica entre as formas verbais singular e plural nas variedades da Língua Portuguesa (cf. Capítulo 4 deste livro).

3.2.3. LEVANTAMENTO E VERIFICAÇÃO DO PERFIL DAS VARIÁVEIS QUE SE MOSTRARAM ESTATISTICAMENTE RELEVANTES

A verificação do perfil das variáveis indicadas como estatisticamente relevantes em estudos de amostras brasileiras e europeias pode fundamentar o reconhecimento dos contextos de concordância variável. Considerando três dos estudos baseados em amostras brasileiras constituídas na segunda metade do século XX, podem-se observar as variáveis linguísticas relevantes em cada nível de escolaridade.

Quadro 4. Variáveis relevantes em estudos da concordância verbal de P6: amostras brasileiras com diferentes níveis de escolaridade

Nível superior de escolaridade (GRACIOSA, 1991)	Nível intermediário de escolaridade (SCHERRE; NARO, 2006)	Ausência de escolaridade (VIEIRA, 1995)
Posição do sujeito em relação ao verbo	Saliência fônica (exceto 9-11 anos)	Saliência fônica
Distância entre sujeito e verbo	Posição do sujeito em relação ao verbo	Paralelismo oracional e discursivo
Paralelismo discursivo	Paralelismo oracional e discursivo	Posição do sujeito em relação ao verbo
[Traço semântico do sujeito – sem dados ou não controlado]	Traço semântico do sujeito	Traço semântico do sujeito
	[Presença do relativo *que*]	Distância entre sujeito e verbo
	Presença de pausa: sujeito e verbo	

O quadro acima aponta as variáveis que se mostraram relevantes em todas as variedades sociais brasileiras: ordem do sujeito em relação ao verbo (que, em alguns estudos, inclui a presença do relativo *que*), distância entre sujeito e verbo (ou variáveis correlatas, como a presença de pausa), paralelismo discursivo, traço semântico do sujeito. Excetuando-se a fala de indivíduos altamente escolarizados, também atuam na amostra brasileira a saliência fônica e o paralelismo oracional. No presente estudo, resguardadas as características do *corpus* – que abrange os níveis de escolaridade fundamental, médio e superior, em conjunto –, a atuação dessas variáveis é confirmada.

É relevante observar que as influências de ordem social – especialmente grau de escolaridade – aceleram ou retraem a implementação das restrições linguísticas. Ao que tudo indica, essa diferença de comportamento verificada nas amostras de indivíduos com curso superior relaciona-se à avaliação da concordância em cada comunidade de fala. Por constituir um estereótipo (LABOV, 1972), a ausência da marca de plural ocorre na fala de indivíduos escolarizados brasileiros somente em construções que podem ser consideradas não marcadas ou pouco salientes, aquelas que só chegam ao nível da consciência em situações mais monitoradas: (i) "sujeito" posposto (*chegou os livros*); (ii) sujeito anteposto, quando distante do verbo, representado pelo relativo *que*, com discreta diferenciação

fônica entre as formas singular e plural (*o rapaz e a moça que sempre de manhã pode ver o sol nascer*). Assim, no PB, constitui traço indexical de iletramento/pouca instrução a ausência da marca de plural em contextos salientes: sujeito anteposto (sobretudo quando não marcado), próximo ao verbo com alto grau de saliência fônica (como em *os menino saiu*).

Os padrões verificados nos dados produzidos por brasileiros com curso superior aproximam-se quantitativa e qualitativamente dos observados nos trabalhos com amostras europeias. O quadro a seguir, que apresenta resultados apenas dos estudos que se valeram da seleção estatística de grupos de fatores possivelmente atuantes no PE, confirma essa semelhança.

Quadro 5. Variáveis relevantes em estudos da concordância verbal de P6 com amostras europeias

Monguilhott (2009)	Rubio (2012)	Monte (2012)
Traço semântico do sujeito	Posição do sujeito em relação ao verbo	Posição do sujeito em relação ao verbo
Posição do sujeito em relação ao verbo	Traço semântico do sujeito	Traço semântico do sujeito
Tipo de verbo	Tipo morfológico do sujeito	Tipo estrutural do sujeito/SN
		Tipo de verbo (verbo 'ser' *versus* outros verbos)
		Saliência fônica

O tratamento estatístico nos trabalhos que adotaram a metodologia variacionista revela que, no PE, atuam sistematicamente os grupos de fatores referentes à posição e ao traço semântico do sujeito. Em segundo plano, é também reveladora a seleção das variáveis tipo de verbo e tipo estrutural do sujeito em dois dos três trabalhos, que, em linhas gerais, permitiram aferir principalmente a influência dos contextos de núcleos coordenados com elementos singulares, com o relativo *que* e, ainda, com o sujeito não expresso. O controle do tipo de verbo, por sua vez, permite validar sobretudo a particularidade das construções copulativas no PE.

Conforme já se afirmou aqui (cf. BRANDÃO; VIEIRA, 2012b; VIEIRA; BRANDÃO, 2014), o estatuto dos grupos de fatores referentes à posição e ao traço semântico do sujeito em nada colabora para a definição das características próprias do PE, do PB ou de qualquer variedade. Na

realidade, trata-se de uma influência de natureza supostamente universal[17], que ocorre em línguas e variedades linguísticas de origens absolutamente distintas, como bem demonstra Corbett (2000):

Tabela 22. Realização da marca de plural verbal de acordo com a ordem e o traço semântico do SN sujeito segundo Corbett (2000)

	Animate		Inanimate	
	N	%PL	N	%PL
Subject-predicate				
Medieval Spanish	288	96	243	31
German	1.095	96	1.702	67
Russian	115	100	67	85
Serbo-Croat	21	100	35	91
Predicate-subject				
Medieval Spanish	318	69	239	6
German	379	93	925	40
Russian	89	84	114	28
Serbo-Croat	23	70	62	26

Fonte: Corbett, 2000, p. 201

Quanto ao grupo saliência fônica, selecionado sistematicamente apenas no trabalho de Monte (2012), entende-se que sua suposta relevância esteja vinculada à superposição com o controle de duas construções: (i) estruturas com verbo *ser* copulativo; (ii) presença de sujeitos pospostos e/ou de estruturas com possíveis interpretações de tópico e sujeito nulo expletivo *isso*.[18] Ademais, destaca-se que a observação das características relativas à pronúncia das desinências número-pessoais, detalhadamente proposta no Capítulo 4 deste livro, reforça que é preciso reconfigurar os níveis de diferenciação fônica entre as formas singular e plural em consideração aos traços fonético-fonológicos presentes nas variedades europeias, como já se propôs anteriormente.

[17] Diversos estudiosos já aventaram a hipótese de universais em variação, seja no âmbito geral, translinguístico, ou particular a cada língua ou tipo linguísticos. Consultar, a esse respeito, Chambers (2004), Szmrecsanyi; Kortmann (2009); Trudgill (2009).

[18] Verificou-se, em Monte (2009), que os únicos três dados sem concordância envolvendo verbos de alta saliência fônica, na rodada sem as ocorrências do verbo *ser*, são dos seguintes tipos: sujeito posposto, em dois casos (*começo os anos a passare*), e sujeito possivelmente interpretado como um expletivo *isso* (*já tem se feito (algumas) greves mas nunca deu... nunca deu resultado...*).

4. CONSIDERAÇÕES FINAIS: A NATUREZA DAS REGRAS DE CONCORDÂNCIA EM VARIEDADES DO PORTUGUÊS

Após a extensa descrição de dados de P6 nas variedades brasileiras, africanas e europeias do Português e a abordagem contrastiva de resultados presentes nesta e em outras investigações, entende-se ser necessário dar relevo às contribuições do trabalho, a fim de postular algumas generalizações e novas questões científicas.

Em função das particularidades características das amostras europeias, que as diferenciam das brasileiras e das africanas, podem-se postular evidências quantitativas e qualitativas suficientes para fundamentar a proposta, genericamente apresentada no Capítulo 1 deste livro, de que existem na língua portuguesa dois padrões de concordância: um semicategórico, típico do PE, de forma mais evidente nas variedades continentais, praticadas por pessoas escolarizadas ou não; e outro, variável, suscetível à ação de restrições linguísticas claras – que fazem com que a alternância concordância e não concordância se aplique à ampla diversidade de estruturas – e extralinguísticas – relacionadas sobretudo ao grau de instrução do indivíduo e ao perfil rural x urbano.

Esse padrão semicategórico – associado às tendências também verificadas no caso da concordância interna ao SN, considerada categórica por Brandão (2013), e da concordância verbal de primeira pessoa plural, em que se verifica a taxa de 100% de expressão plural com a forma *nós* (VIANNA, 2011; RUBIO, 2012; VIEIRA; BRANDÃO, 2014) – também é verificado na fala culta brasileira (assim como na são-tomense, conforme BRANDÃO; VIEIRA, 2012a), para a qual a não realização da marca plural constitui um verdadeiro estereótipo. Na fala de indivíduos com níveis baixo ou médio de escolaridade, especialmente dos meios rurais, o padrão variável de concordância é flagrantemente observado. Nesse sentido, pode-se postular, à semelhança do que faz Scherre; Naro (2006), a convivência de uma norma de concordância prestigiosa, semicategórica *versus* uma regra variável natural na realidade brasileira.

Em São Tomé, ao mesmo tempo em que se verifica a atuação de uma regra variável, com forte preferência pela concordância, chama a atenção o fato de que os dados são claramente influenciados por condicionamentos sociais. Ser escolarizado promove claramente o domínio da língua portuguesa e, por consequência, o da concordância. De todo modo, a presença de uma regra variável em variados contextos estruturais soma-se às evidências relativas à perda de marcas flexionais típicas da situação de intenso contato linguístico.

Do exposto, é inevitável concluir que o PE (e o PB falado por pessoas altamente escolarizadas) não partilha o mesmo perfil quantitativo nem qualitativo do PB vernacular. Restrições sociais, como escolaridade, perfil urbano/rural, sexo do informante, e a presença ou a ausência de formas não marcadas em estruturas morfossintáticas específicas (verbo com alta saliência fônica na oposição singular/ plural; ordem direta; sem elemento interveniente entre referente sujeito e verbo) permitem estabelecer condições exclusivamente brasileiras, que permitem estabelecer semelhanças ou diferenças sociais/regionais e identificar os falantes.

De outro lado, não se encontram evidências que permitam postular o valor social da variação de concordância no PE, sobretudo na amostra continental[19]. Em outras palavras, os contextos em que se constata a suposta falta de concordância nos dados europeus, de caráter universal, nada podem revelar do estatuto social da variedade em questão. A relevância de restrições linguísticas gerais – como a posição do sujeito, traço semântico do sujeito ou até paralelismo discursivo – nada pode dizer sobre a origem histórica do PB e do PE. Antes, constituem tão somente o registro do comportamento específico de estruturas sintática, semântica e foneticamente particulares; de fato, o conjunto de fatores em atuação nas três variedades do PE parece obedecer a condicionamentos morfofonológicos (sândi externo) e sintático-semânticos do tipo genérico ou "universal", não podendo ser indicadores que caracterizam cada uma destas variedades do PE (BRANDÃO; VIEIRA, 2012b).

Nesse sentido, os dados apresentados neste estudo não apresentam evidências que sustentem a proposta de que teria havido, na constituição do PB, a expansão das estruturas europeias. Ao menos em dados contemporâneos, não parece justificável imaginar que, de contextos tão particulares (como, por exemplo, os de sujeito posposto, construções particulares com verbo *ser*, ou ainda da presença de SN topicalizados e sujeito nulo expletivo), tenham emergido os padrões variáveis brasileiros de concordância. Ainda que a saliência fônica atuasse nos dados europeus, não parece sustentável propor que de uma ocorrência como "eles canta" derivasse uma estrutura como "nós sai". Apelar para a analogia (como faz RUBIO, 2012), nesse caso, parece fazer retroagir a um tempo em que as teorias e os métodos científicos eram insuficientes para a explicação dos dados linguísticos. Na opinião de Lucchesi (2012, p. 263), mais do que isso, "invocar a analogia, tão cara à Filologia do século XIX, nesse caso, não é apenas um anacronismo, é um desatino teórico e epistemológico, até porque, assim é possível explicar qualquer coisa (...)".

[19] Para maiores detalhes sobre o comportamento do PE insular, vale a pena consultar o próximo capítulo deste livro.

Considerando o Princípio do Uniformitarismo, os diferentes padrões de concordância favorecem a tese de diferentes processos históricos de formação, segundo a qual o PB, dada a extensão dos contextos sem desinências plurais e a configuração de normas sociais diferenciadas, à semelhança do que ocorre hoje no PST, reflete claramente o perfil polarizado dos contatos linguísticos, que opunha e opõe ainda hoje os meios rural e urbano.

Por fim, espera-se que a descrição da expressiva quantidade de dados, coletados no âmbito do Projeto *Estudo comparado dos padrões de concordância em variedades africanas, brasileiras e europeias*, possa ter colaborado para a resolução do problema de validação empírica das propostas interpretativas existentes, tendo metonimicamente contribuído para o entendimento de questões diversas atinentes à relação entre a variação linguística e a caracterização dos dialetos sociais. Embora a interpretação relativa à história da constituição das línguas e suas variedades permaneça um estimulante desafio aos estudos linguísticos, ainda com questões a responder, espera-se que o presente estudo tenha provido informações relevantes para a comprovação científica das hipóteses formuladas. Sem dúvida, a conjugação desses resultados aos obtidos em outras estruturas, como no caso da concordância nominal e verbal de primeira pessoa plural (cf. VIEIRA; BRANDÃO 2014), amadurecerá a proposta ora apresentada.